Solving problems of agricultural production development by modeling using the cluster analysis method

Tatiana Lubova*, *Guzel* Salimova, *Gulnara* Nigmatullina, *Damir* Islamgulov, and *Irek* Araslanbayev

Bashkir State Agrarian University, 34, 50th anniversary of October street, Ufa, 450001, Republic of Bashkortostan, Russian Federation

Abstract. The article outlines the need to apply a systematic approach to the analysis and solution of agricultural production development problems. The system approach implementation is carried out using a multidimensional data analysis method – the cluster analysis method. A system of indicators characterizing the conditions and results of agricultural enterprises in the regions of the Central and Volga Federal Districts of the Russian Federation has been built, a multidimensional grouping of regions has been carried out to identify problems and directions for their solution. The multidimensional approach to the study allowed an objective analysis and formulation of well-founded specific conclusions and recommendations for the regions of each cluster. As a result, two clusters of regions were identified according to the level of agricultural production development. Subclusters are also allocated within clusters to detail and concretize proposals. Recommendations on agricultural production development are made to the regions of each cluster. The directions for the formation of the strategy for the development of agriculture of the corresponding groups of regions and territories are outlined.

1 Introduction

Agricultural production is strategically important for the country and its regions, for the well-being and health of the population, it ensures food security and, to a large extent, the independence of the state. Scientists in their research address various aspects of the current stage of agricultural production functioning and development. We have approached the analysis of modern research in the field of agricultural production and agriculture based on two our interests: to study the directions of research on agricultural development problems, as well as to analyze studies that use the cluster analysis method.

The relevance of the study of the well-being of agricultural territories depending on the economic activity of economic entities on their territory is confirmed by Sanjib Mondal, Pritam Ghosh, Pratima Rohatgi. Understanding of the spatial model of livelihood provision is important for poverty reduction and eradication of spatial inequality in social and economic

* Corresponding author: lubova@list.ru

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

conditions [1]. C. Castillo et al. note that urbanization, demographic changes, transition to digital and energy technologies, climate change, and increasing inequality affect the development of territory and agricultural production [2]. Scientists widely address the issues of industry digitalization. Jakku, E. et al. conclude that digital transformation in agriculture includes advances in information and communication technologies, it will contribute to increased productivity and efficiency while reducing risks and negative consequences [3]. Xia, J.at the same time notes that the introduction of digital technologies in agricultural production should be carried out considering the social and economic, technological, and institutional features of the state and territories of the point of view [4]. One of the directions of agricultural production digitalization is the use of digital technologies in the analysis of data on the industry development, identifying problems and developing ways to their solution. Pagliacci, F., Salpina, D. apply the cluster analysis method to identify territorial foci of geographical indications of the agri-food industry at risk of climatic disasters. They perform cluster data analysis at the municipality level. The analysis covers physical, social, and production characteristics [5]. Yi, J., Fang, Z., Yang, G., He, S., Gao, S. note that cluster analysis, as one of the main methods of data mining, it is important for discovering the natural structure of data to obtain useful information from huge data volumes. Nevertheless, they note that many existing and applied clustering algorithms have the problem of low clustering accuracy, they offer updated, improved clustering algorithms [6]. Bartesaghi, P., Clemente, G.P., Grassi, R., Luu, D.T. apply cluster analysis in the study of trading operations and note that clusters can change, develop over time. In general, not only their internal composition is changing, but also the ranking of the central members within. These changes in large international clusters as well as regional ones will reflect the results and dynamics of cooperation, choice of partners, and competition between industries and countries [7]. We believe that this method advantage is especially important in modern conditions of changing the structure of the economy, areas of cooperation of territories, countries. Xuesong Gao, Lun Liu, Dinghua Ou, Haomiao Yuwu combine the traditional method of cluster analysis and modern digital data analysis technologies. They use big mobile phone data, which is still new to rural research, to study the spatial model and factors affecting mobility in rural areas [8]. Thus, traditional, familiar, well-established methods of cluster analysis are being modernized considering new technologies for data collection and analysis [15].

Thus, the issues of agricultural production development, application of research methods that allow a systematic and multilateral approach to its study, are relevant and widely discussed by scientists. The purpose of this study is to build models that allow to identify the problems of agricultural production development and identify ways of their solution. In accordance with the purpose, the tasks are defined:

- to build a system of indicators for grouping regions by the method of multidimensional analysis;
- to identify groups and subgroups (clusters and subclusters) of regions that are united and similar in terms of the indicators under consideration;
- to characterize the selected clusters by the average values of the system indicators;
- to develop recommendations to the regions of each cluster and subcluster to solve the problems of agricultural production development, develop strategies for its development for certain territories;
- to determine the vector of further research of agricultural production, research methodology.

2 Materials and Methods

The research methodology is presented by the cluster multidimensional analysis method. As the main advantage of the cluster analysis grouping method, the possibility is noted for

grouping units of the studied population not by one trait or indicator, but simultaneously by the entire set of features. At the same time, the method of cluster analysis implementation does not limit the type of objects under consideration, it studies a lot of presented empirical data of any dimension and of any kind.

The sources of information for analysis were an array of data provided by the official statistical service of Russia, presented in the Rosstat statistical collection [9]. From the calculated and published statistical indicators, 16 indicators were selected that characterize the resource provision and efficiency of agricultural enterprises [16]. A system of indicators for analysis is constructed (Table 1).

Table 1. System of indicators characterizing the activities of agricultural enterprises.

Name of indicators
Availability of labor, people per 1000 hectares of farmland
Share of fixed assets per 1 person employed in agriculture, thousand rubles (capital-labor ratio)
Share of fixed assets per 1 hectare of farmland, thousand rubles (capital-area ratio of farmland)
Mineral fertilizers were applied to 1 hectare of sown area of agricultural crops, kg of active substance
Feed consumption per 1 conditional head of cattle, hundredweight of feed units
Depreciation of fixed assets of agricultural industries, as a % of the value of the initial cost of fixed assets
Value of investments in fixed assets, thousand rubles per 100 hectares of farmland
Yield of grain and leguminous crops, centners per 1 hectare
Potato yield, hundredweight per 1 hectare
Milk yield per 1 cow in agricultural organizations, kg
Value of gross agricultural output, thousand rubles per 1 employed person (labor productivity indicator)
Value of gross agricultural output per 1 hectare of farmland, thousand rubles
Value of gross output in agriculture, rubles of output per 1 ruble of the value of fixed assets (capital productivity ratio indicator)
Financial result in agriculture per 1 enterprise, million rubles
Sale profitability of crop products, %
Sale profitability of livestock products, %

The STATISTICA data processing program was used for cluster analysis. Instead of the initial values of the variables included in the analysis when grouping regions, we use standardized variables obtained as a result of the normalized transformation. Standardization makes it possible to bring indicators measured in different scales and expressed in different units of measurement into a single form. For the purpose of visual representation of the formed clusters, a graphical method is used – a cartogram of the placement of region clusters on the territory is constructed.

3 Results

As a result of the cluster analysis, the regions of the Central and Volga Federal Districts were divided into 2 clusters. The unification dendrogram is shown in Figure 1. In turn, subclusters can also be allocated as part of the first cluster. Subclusters are allocated to specify the planned measures to increase the efficiency of agricultural production.

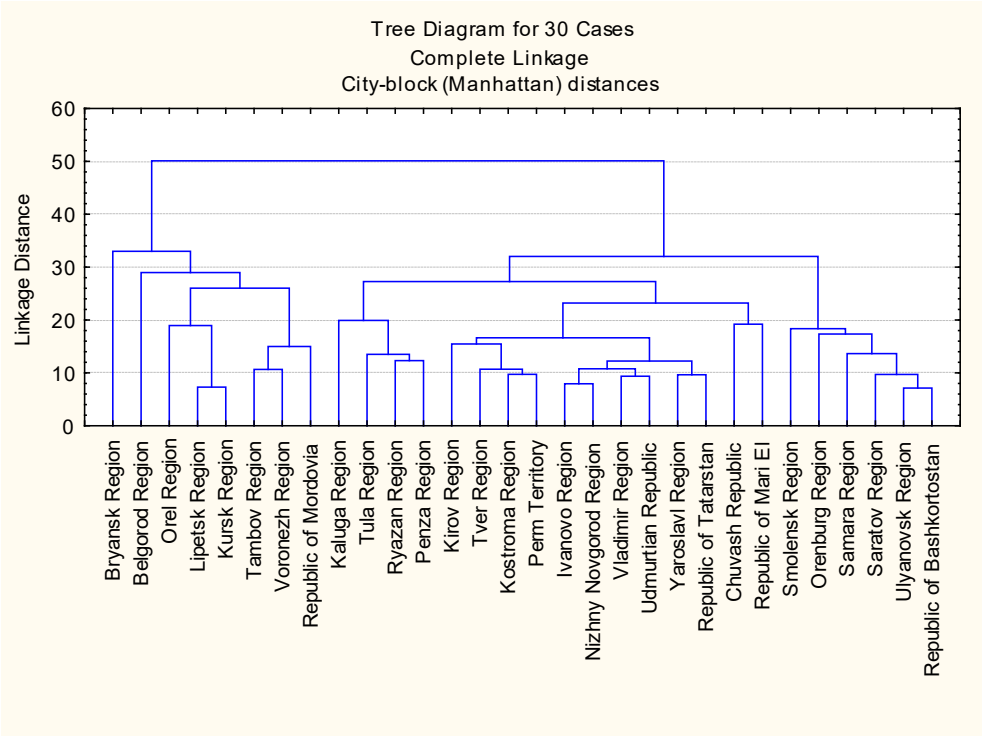

Fig **1.** Dendrogram of the regions of the Central and Volga Federal Districts.

The selected groups of regions are presented in Table 2.

Table2. Clusters of regions of the Central and Volga Federal Districts of the Russian Federation.

Cluster	Subcluster	Cluster regions (the numbers of regions on the cartogram are shown in brackets)
I	1	Republics: Bashkortostan (1) Regions: Ulyanovsk (3), Saratov (2), Samara (5), Orenburg (4), Smolensk (6)
I	2	Republics: Mari El (12), Chuvash (18)
I	3	Republics: Tatarstan (14), Udmurt (15), Regions: Yaroslavl (17), Vladimir (16), Nizhny Novgorod (10), Ivanovo (8)
I	4	Perm Territory (7) Regions: Kostroma (11), Tver (9), Kirov (19)

I	5	Regions: Penza (21), Ryazan (23), Tula (24), Kaluga (20)
II	-	Regions: Bryansk (29), Belgorod (30), Oryol (28), Lipetsk (26), Kursk (25), Tambov (27), Voronezh (22), Republics: Mordovia (13)

* the numbers of the regions on the cartogram are shown in brackets (Figure 3).

The distribution of regions by clusters is clearly shown in Figure 3.

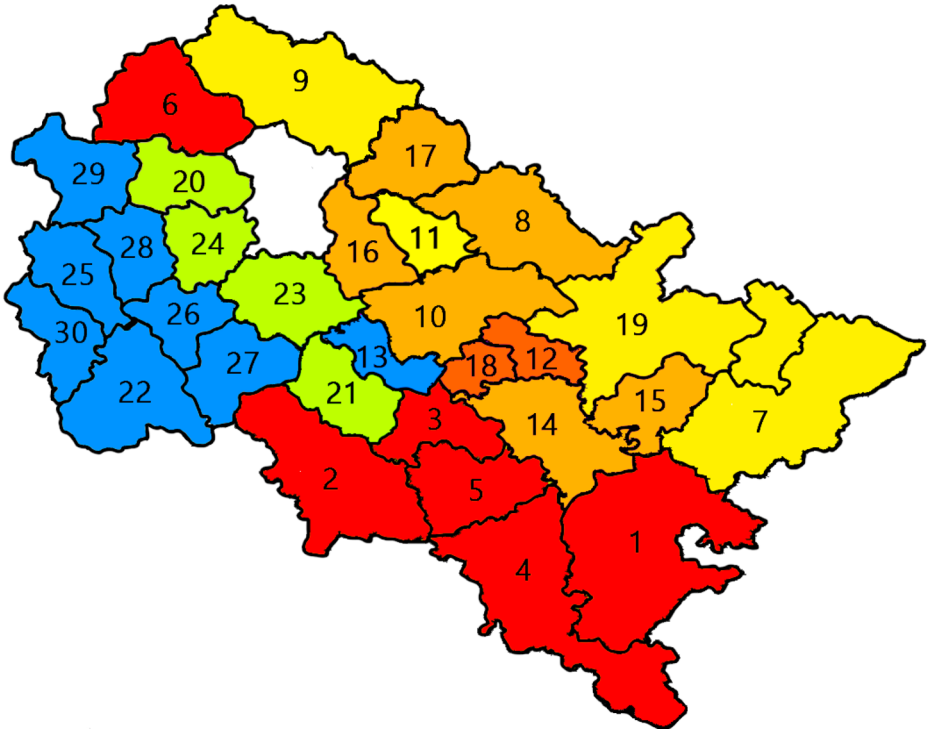

Fig **2.** Cartogram of the distribution of the regions of the Central and Volga Federal Districts by clusters.

Cluster 1 includes 22 regions (73.3% of the total analyzed population). Cluster 2 united 8 regions (26.7% of the analyzed total). Let's characterize each cluster by the average values of the system indicators (Table 3).

Table 3. Values of indicators in clusters of regions of the Central and Volga Federal Districts of the Russian Federation.

Indicators	Cluster 1	Cluster 2	On average in the Russian Federation
Number of regions, units	22	8	×
Potato yield, hundredweight per 1 hectare	158.9	152.1	160.0

Yield of grain and leguminous crops, centners per 1 hectare	18.6	38.6	26.7
Mineral fertilizers were applied to 1 hectare of sown area of agricultural crops, kg of active substance	49.6	132.3	74.6
Milk yield per 1 cow in agricultural organizations, kg	6870	7411	7007
Feed consumption per 1 conditional head of cattle, hundredweight of feed units	32.8	24.9	28.6
Depreciation of fixed assets of agricultural industries, as a % of the value of the initial cost of fixed assets	43.1	43.5	43.2
Share of fixed assets per 1 hectare of farmland, thousand rubles (capital-area ratio of farmland)	37.0	83.6	4.68
Share of fixed assets per 1 person employed in agriculture, thousand rubles (capital-labor ratio)	1906	3054	1783
Value of gross output in agriculture, rubles of output per 1 ruble of the value of fixed assets (capital productivity ratio indicator)	1.02	1.03	0.96
Value of gross agricultural output per 1 hectare of farmland, thousand rubles	34.55	85.19	4.50
Availability of labor, people per 1000 hectares of farmland	20.19	30.49	2.62
Value of investments in fixed assets, thousand rubles per 100 hectares of farmland	281.2	862.6	36.7
Value of gross agricultural output, thousand rubles per 1 employed person (labor productivity indicator)	1812	3013	1717
Financial result in agriculture per 1 enterprise, million rubles	2.98	45.19	8.74
Sale profitability of crop products, %	19.71	69.45	48.50
Sale profitability of livestock products, %	9.81	17.68	12.60

The data in Table 3 allow to conclude that the second cluster includes regions that are characterized by the most favorable natural and climatic conditions for both crop production and livestock production. In addition, the enterprises of the second cluster are the most provided with resources, including labor, land resources, technical equipment, scientific achievements and developments. This made it possible to achieve higher agricultural production results.

The first cluster turned out to be numerous, and, in turn, its regions were also divided into 5 subclusters by cluster analysis. Subcluster 5 includes 4 regions with the maximum level of investment in fixed assets per 100 hectares of farmland, which has significantly improved the material and technical base and introduced advanced agricultural production technologies. But at the same time, the low availability of labor did not allow the regions under study to enter the second cluster. It is recommended for these regions to pay more attention to the attractiveness of the territory for population living and employment.

Proximity to the Moscow region draws the able-bodied population to work, forming a commutation.

Two regions with the loss-making capacity of the crop industry and, accordingly, the minimum size of the financial result per agricultural enterprise are allocated to a separate subcluster. The regions of this subcluster are characterized by animal husbandry. Nevertheless, it is also recommended for the regions of this subcluster to pay attention to the channels for the sale of crop products, their quality, marketing policy, and marketing activities.

The main limiting factor of the first subcluster regions is the size of the territory and its low availability of resources. The available resources are used efficiently, but their insufficiency hinders the development of agricultural production. It is recommended to direct scientific research to modeling optimal transportation plans, logistics solutions, and transport industry development. We propose to form a management center for rural territories not only on the scale of one region, but to apply methods of cooperation and industrial integration.

4 Discussion

The results of our research are consistent with the directions and results of other scientists studying rural areas and agricultural production, research methods. Kulshrestha, S.K. notes that an integrated, collaborative and proactive approach to regional spatial planning is needed, which promotes inclusive and sustainable development, capable of adapting to technological innovations and climate change [10]. Scientists combine traditional and new knowledge, which, in our opinion, should be carried out when determining the directions and potential for agricultural production development. We come to the conclusion about the importance of the management organization for the sustainable development of agricultural production, which is confirmed by the results and conclusions of other scientists. C. Castillo et al. note that the development of spatial planning considering the biophysical, social, economic, and cultural characteristics of regions can help to realize the potential of rural areas [2]. We consider it necessary to strengthen ties between the regions, to create a single management center for the development of rural areas and agricultural production, in particular, within the framework of the obtained clusters. Cavallaro, F., Dianin, A. note that the main attention in the organization of production in rural areas is paid to transport [11]. Chan, B.R.A. et al. say that rural roads mean a lot for the economic and social development of society and the rural territories [12]. C. Castillo et al. state that a comprehensive sustainable territorial development is necessary, considering future technological changes (innovations, digitalization, transport systems, housing, migrants, etc.) [2]. In our study, we also came to the conclusion that, in particular, for the regions of one of the subclusters, it is necessary to develop transport infrastructure for agricultural production development. Miroro, O.O., et al. consider factors affecting farmers' decisions to join agricultural cooperatives [13]. We are also talking about the need to develop a system of cooperation.

C. Castillo, et al. note the importance of applying analytical methods and forecasting methods to effectively assess the problems that agricultural production faces [2]. We develop the methodology of analysis and apply a multidimensional approach to the analysis of agricultural production development. The need to apply a whole system of indicators for the development of agricultural production and solving emerging problems, justified in our study, is also confirmed by scientists. S. Mondal, P. Ghosh, P. Rohatgi assessed the livelihood of rural residents by combining the availability of natural resources, housing, drinking water and sanitary facilities, economic status, socio-cultural status and the availability of assets. Differentiation of rural settlements according to the level of provision with natural, labor, and technological resources can help politicians develop adequate, sustainable policies at the micro level that increase social and economic sustainability and security of livelihoods, as

well as reduce spatial inequality [1]. In the process of allocating clusters, we also noted that the availability, provision of agricultural production, rural territories with resources, their rational use is a determining factor for the effective development of agricultural production, therefore, for the well-being of the territory. We also agree that the allocation of clusters and their analysis allows to conduct a scientifically sound policy for the development of production and regions in general. The development for our research is that the indicators can be presented in a more expanded form, with justification for their implementation in international studies.

Zeidmane, A., Vintere, A.A. note that new digital technologies are changing entrepreneurship. Entrepreneurship in the digital environment includes innovative ways and channels of finding customers and business partners, the development of new innovative products, new sources of income, new options for cooperation with platforms and partners. The modern labor market requires employees to possess basic digital skills, general, special, and additional skills in the field of information and communication technologies [14]. We support these conclusions, and in our study we also talk about the development and application of digital technologies in agriculture, therefore, about the need for the formation of digital competencies among graduates of agricultural specialties in educational institutions of vocational education. This will contribute to the production development, solving the problem of ensuring the quality of the workforce, timely detection of development problems and their prompt solution. In the course of the study, we were also interested in how the region ended up in a certain cluster, what is the forecast of indicators and scenarios for the development of this cluster distribution, which will be the reason for the transition of the region from one cluster to another in dynamics. These challenges expand the scope of our intended research and will lead to meaningful results.

5 Conclusion

The results of the analysis show that the regions of the second cluster are characterized by the most favorable natural and climatic conditions for the production of agricultural products. The enterprises of the second cluster are the most provided with resources, which made it possible to achieve higher results of agricultural production. The regions of the first cluster, depending on their location in a particular subcluster, are recommended to pay more attention to the attractiveness of the territory for living and employment of the population, to pay attention to the channels of sale of agricultural products, their quality, marketing policy, marketing activities. It is recommended to direct scientific research to modeling optimal transportation plans, logistics solutions, and transport industry development. We propose to form a management center for rural territories not only on the scale of one region, but to apply methods of cooperation and industrial integration.

Further research will be aimed at developing strategies for agriculture development in the regions of each of the clusters and subclusters, considering dynamic changes, the introduction of digital technologies in the industry. The research methodology will be expanded by the use of panel data analysis methods, which will allow to identify the main factors and dynamic dependencies in spatio-temporal data.

References

1. S. Mondal, P. Ghosh, P. Rohatgi, Village-level livelihood security: A case study on a wasteland-dominated forest fringe region of rural India (2022) https://doi.org/10.1111/rsp3.12552

2. C. P. Castillo, S.van Heerden, R. Barranco, C. Jacobs-Crisioni, M. Kompil, A. Kučas, J. P. Aurambout, F. Batista e Silva, C. Lavalle. Urban–rural continuum: an overview of their interactions and territorial disparities (2022) https://doi.org/10.1111/rsp3.12592
3. E. Jakku, S. Fielke, A. Fleming, C. Stitzlein, Sociologia Ruralis, **62(2),** 363-388 (2022)
4. J. Xia, Telecommunications Policy, **46(2),** 102258 (2022)
5. F. Pagliacci, D. Salpina, Land Use Policy, **123,**106404 (2022)
6. J. Yi, Z. Fang, G. Yang, S. He, S. Gao, Knowledge-Based Systems, **258,** 110004 (2022)
7. P. Bartesaghi, G.P. Clemente, R. Grassi, D.T. Luu, Journal of Economic Behavior and Organization, **204,** 304-341 (2022)
8. X. Gao, L. Liu, D. Ou, H. Yuwu, Using mobile phone data to probe the mobility-related well-being of rural residents: A case study of Chengdu, China. (2022) https://doi.org/10.1111/rsp3. 12604
9. *Regions of Russia. Socio-economic indicators* (Moscow: Rosstat, 2022) 1122.
10. S.K. Kulshrestha, Emerging Advances in Regional Spatial Planning. Advances in 21st Century Human Settlements, 147-175 (2022)
11. F. Cavallaro, A. Dianin, Journal of Transport Geography, **98** 103224 (2022)
12. B.R.A. Chan, B.M. Nazri, O.K.R.R.A. Abdullah. Journal of Sustainability Science and Management, **17(6),** 24-37(2022)
13. O.O. Miroro, D.N. Anyona, I. Nyamongo, (...), K. Waweru, L.Kiganane. *Determinants of smallholder farmers' membership in co-operative societies: evidence from rural Kenya.* International Journal of Social Economics (2022).
14. A. Zeidmane, A.*A. Vintere, Case study of students' views on the digital skills needed for the labour market. 18th International Conference on Cognition and Exploratory Learning in Digital Age*, CELDA, 163 -169. (2021).
15. G. Salimova, A. Ableeva, T. Lubova, G. Habirov, A. Sharafutdinov, Montenegrin Journal of Economics, **17(4)** 111-131 (2021).
16. A. Ableeva, G. Salimova, R. Gusmanov, T. Lubova, O. Efimov, A. Farrahetdinova Montenegrin Journal of Economics, **4(15),** 183-193 (2019)